

Unit 5 Project Assessment

Mark Schulist

2023-02-17

How Quickly Do Synthetic Fabrics Such as Polyester Decay in Landfills?

We want to know if the data give convincing evidence that polyester decays more in 16 weeks than 2 weeks.

Design

This is an experiment because the 10 strips were randomly assigned to two groups—buried for 2 weeks or buried for 16 weeks.

The observational units are the polyester strips. There are 10 of them.

They randomly assigned each strip to be left for 2 or 16 weeks.

Explanatory variable: number of days the polyester strip was left in landfill (qualitative—two discrete groups)

Response variable: breaking strength of each polyester strip after being left in the landfill

The samples are independent. They are not interacting with each other and hopefully randomized enough. We should use the multiple means procedure.

The parameter of interest is the difference in means (μ) between leaving the strip for 2 weeks and 16 weeks.

H_0 : There is no difference in the amount of strength required to break polyester between being in a landfill for 2 weeks and being in a landfill for 16 weeks.

H_a : Polyester in a landfill for 16 weeks requires less strength to break than polyester in a landfill for 2 weeks.

$$H_0 : \mu_{2 \text{ weeks}} - \mu_{16 \text{ weeks}} = 0$$

$$H_a : \mu_{16 \text{ weeks}} > \mu_{2 \text{ weeks}}$$

Data Exploration

Group 1 (2 weeks)	118	126	126	120	129
Group 2 (16 weeks)	124	98	110	140	110

The sample size is 10 since there are 10 strips of polyester that were put into the landfill.

```
two_weeks <- c(118, 126, 126, 120, 129)
sixteen_weeks <- c(124, 98, 110, 140, 110)
all_data <- c(two_weeks, sixteen_weeks)
sample_stat <- mean(two_weeks) - mean(sixteen_weeks)
```

The sample statistic (μ_{diff}) is 7.4 (123.8 - 116.4).

Inferences

In order to conduct a simulation, we would have 10 cards with each of them containing the breaking strengths of one strip of polyester. We would shuffle the cards and deal them into two piles of five. One pile would represent the two week group (group 1) and the other would represent the sixteen week group (group 2). We would then record the difference in means of each of the groups and add it to a dot plot (or histogram). We would repeat the shuffling many (thousands) of times and then compare our sample statistic to our distribution.

```
shuffle <- function(all_data) {
  shuffled <- sample(all_data)
  pile_1 <- shuffled[1:5]
  pile_2 <- shuffled[6:10]
  diff <- mean(pile_1) - mean(pile_2)
  diff
}
```

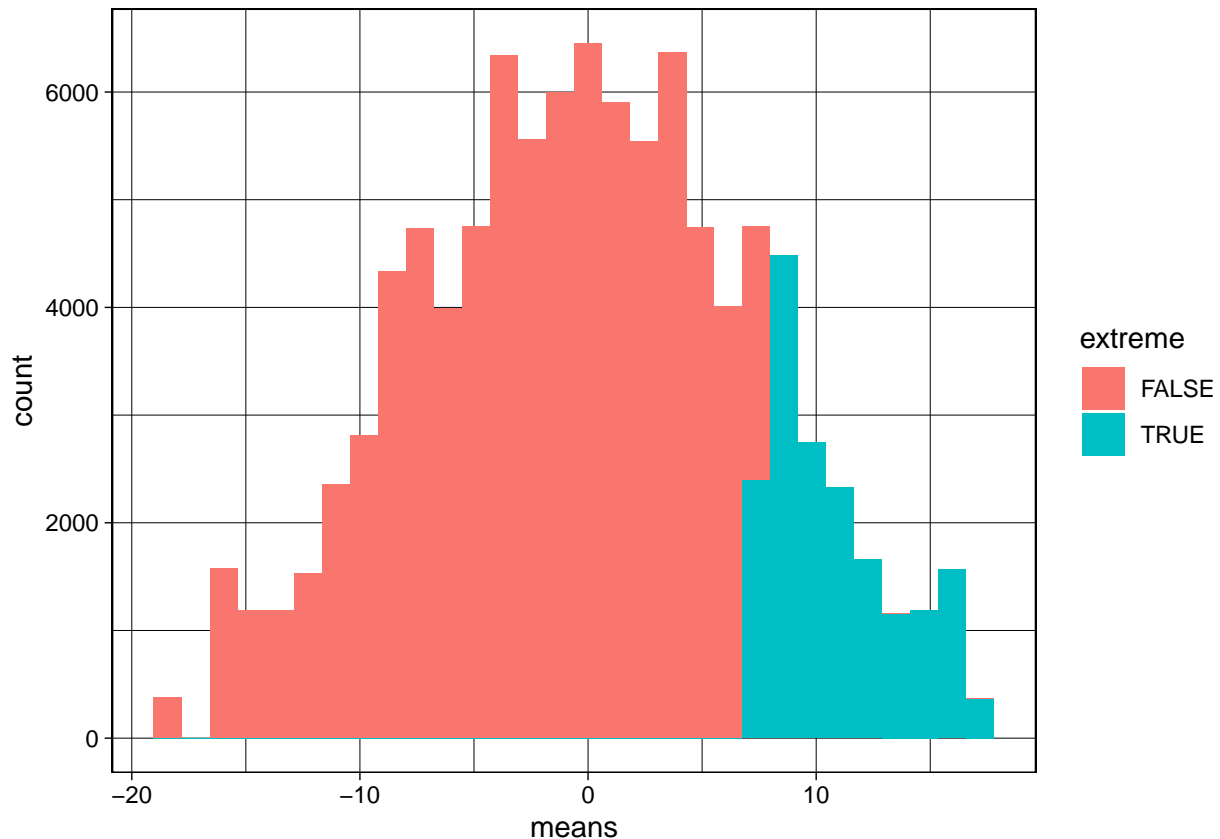
Let's conduct this simulation 100,000 times (because why not).

```
set.seed(123)
means <- replicate(100000, shuffle(all_data))
```

Now we can plot it and see where our observed statistic falls.

```
color_means <- means %>%
  as_tibble() %>%
  rename(means = 1) %>%
  mutate(extreme = means > sample_stat)

ggplot(color_means, aes(means, fill = extreme)) +
  geom_histogram(bins = 30) +
  theme_linedraw()
```



So, just from looking at it, it doesn't seem like our statistic is significant, but let's find an "official" p-value just to be sure.

```
p_value <- mean(color_means$extreme)
```

So our p-value is 0.1788. If the amount of time polyester spent in a landfill (2 vs 16 weeks) does not affect its strength, then there is a 17.88% chance of observing a difference of 7.4 or more extreme.

```
sim_mean <- mean(color_means$means)
sim_sd <- sd(color_means$means)
t_score <- (sample_stat - sim_mean) / sim_sd
```

Our t-score is 0.992. This is the number of standard deviations our observed statistic is above the mean of the null distribution (a t distribution).

```
lower <- sample_stat - 2 * sim_sd
upper <- sample_stat + 2 * sim_sd
```

We are 95% confident that the true mean difference in strengths between polyester left in the landfill for 2 weeks and 16 weeks is between (-7.5082, 22.3082).

We can use a t test to find the theory-based t-score and p-value.

```
t <- t.test(two_weeks, sixteen_weeks, alternative = "greater")
t
```

```
##
## Welch Two Sample t-test
##
## data: two_weeks and sixteen_weeks
```

```
## t = 0.98887, df = 4.651, p-value = 0.1857
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -7.932851      Inf
## sample estimates:
## mean of x mean of y
##    123.8    116.4
```

The p-value and t-score created through the theory-based test are very similar to the p-value and t-score created through the simulation. The 95% confidence interval is slightly wider, but it still includes 0. I would use the simulation since we do not have at least 10 samples from each group, and it is hard to tell if the data are normally distributed with so few samples.

Conclusions

Because our p-value is significantly above 0.05 (p-value = 0.179), there is not strong evidence to suggest that polyester placed in a landfill for sixteen weeks degrades in strength more than polyester placed in a landfill for two weeks. This was an experiment, so if there was significant evidence, we would be able to make a cause-and-effect relationship.

Look Back and Ahead

I wonder if different types of polyester would respond differently since there are many types of polyesters that are widely used. I also wonder if landfills that contained more hazardous substances would cause the polyester to decay more, and therefore show a significant difference between the two groups. I would like to see them repeat this experiment except do it in the winter when there is more moisture that would decay the polyester. Why didn't this study use more strips of polyester? They knew that 10 total strips is probably not enough for statistical significance.

Does polyester decay more in landfills in the summer or the winter?